

VU Research Portal

I don't know. The effect of question polarity on no-opinion answers

Kamoen, N.; van de Pol, Jasper; Krouwel, A.P.M.; de Vreese, Claes; Holleman, Bregje

2018

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Kamoen, N. (Author), van de Pol, J. (Author), Krouwel, A. P. M. (Author), de Vreese, C. (Author), & Holleman, B. (Author). (2018). I don't know. The effect of question polarity on no-opinion answers. Online publication or Website, American Association for Public Opinion Research.. <http://www.surveypactice.org/article/3222-i-don-t-know-the-effect-of-question-polarity-on-no-opinion-answers>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

I Don't Know. The Effect of Question Polarity on No-opinion Answers

Naomi Kamoen¹, Jasper Van de Pol², André Krouwel³,
Claes De Vreese⁴, Bregje Holleman⁵

Survey Practice

10.29115/SP-2018-0017

Mar 12, 2018

Tags: question polarity, no-opinion, vaas

¹ Institution: Tilburg University

² Institution: University of Amsterdam

³ Institution: VU University

⁴ Institution: University of Amsterdam

⁵ Institution: Utrecht University

Abstract

A new type of political attitude survey that has gained popularity in Europe and in the United States is the voting advice application (VAA). VAAs provide users with a voting advice based on their answers to a set of attitude questions. In the calculation of this advice, no-opinion answers are excluded. We tested the hypothesis that negative VAA questions lead to more no-opinion answers than their positive equivalents. In a field experiment, visitors (N=41,505) of a VAA developed for the municipality of Utrecht in the Netherlands, were randomly guided to one of the versions of the tool in which the polarity of 16 questions was manipulated. Results do not show an overall effect of question polarity. This overall null finding appears to be caused by contrasting effects for two subtypes of negative questions: Explicit negatives (e.g. not allow) yield more no-opinion answers than their positive counterparts (e.g. allow) do, while the reverse holds for implicit negatives (e.g. forbid).

Introduction

Research from the 1940s onward has shown that respondents more often answer “no” or “disagree” to negative questions than “yes” or “agree” to positive ones (e.g., Rugg 1941). This holds both for questions with an explicit sentence negation such as *not* (“The government cannot cut down on social work”; e.g., Holleman et al. 2016), and for questions with an implicit negation, containing a word with negative valence (cf. Warriner et al. 2013), such as *forbid* (“Do you think the government should forbid the showing of X-rated movies”; e.g., Schuman and Presser 1981/1996). Hence, someone’s opinion about an attitude object seems to be more positive when the question is phrased negatively (for an overview, see Kamoen et al. 2013).

These question polarity effects on the mean substantive answers have sparked a debate on which question wording is best (e.g., Chessa and Holleman 2007; Holleman 2006; also see discussions about unipolar versus bipolar questions, e.g., Friborg et al. 2006; Saris et al. 2010). No-opinion answers are an important proxy for (a lack of) data quality, as survey respondents frequently choose such answers to indicate comprehension problems (e.g., Deutskens et al. 2004; Kamoen and Holleman 2017). The current research therefore investigates the effect of question polarity on *nonsubstantive* answers: the proportion of no-opinion answers. To the best of our knowledge, no-opinion answers have not yet been analyzed as a dependent variable in polarity research before. This is probably because survey respondents shy away from providing no-opinion answers (Krosnick and Presser 2010), which means that a large sample size is needed for demonstrating any effect.

The Complexity of Positive vs. Negative Questions

Survey handbooks acknowledge the advantages of mixing positive and negative wording in sets of questions in order to “alert inattentive respondents that item content varies” (Swain et al. 2008, 116) and also to detect straightliners (e.g., Sudman and Bradburn 1982; Weisberg 2005). Yet, they also warn against using negative questions in abundance (e.g., Dijkstra and Smit 1999; Dillman et al. 2009; Korzilius 2000). This is because negative questions are more difficult to process than their positive counterparts. Outside of a survey context, it has been shown repeatedly that negatives take more processing effort than their positive equivalents (e.g., Clark 1976; Hoosain 1973; Sherman 1973). Horn (1989, 168) summarizes that: “all things being equal, a negative sentence takes longer to

process and is less accurately recalled and evaluated relative to a fixed state of affairs than the corresponding positive sentence”. This holds both for sentences that include an explicit negation (e.g., not happy/happy) as well as for sentences including an implicit negative, and generalizes across morphologically markedness (e.g., unhappy/happy vs. sad/happy), and across semantic types such as verbs (e.g., forget/remember) and contradictory adjectives (e.g., absent/present) (Clark and Clark 1977). The presumed cause for these processing differences is that negatives must be converted into positives before they can be understood (see Clark 1976; Kaup et al. 2006).

A second reason for survey handbooks to advise against the use of negative questions is that the answers to negative questions are relatively hard to interpret. This is because it is counterintuitive for respondents to answer ‘no’ or ‘disagree’ to indicate that they favor an attitude object (Dillman et al. 2009). For questions with an explicit negative, this confusion is particularly large, because in daily language use a *no*-answer to a question with an explicit negation indicates agreement with the negated statement. For example, one would probably answer *No*, *asylum seekers should not be allowed* to indicate agreement with the statement *Do you think the government should not allow any more asylum seekers* (example taken from Dijkstra and Smit 1999, 84). In a survey context, however, a *yes*-answer is the desired response to indicate agreement. This causes difficulties in interpreting the meaning of *yes/no* and *agree/disagree*-answers to questions with an explicit negation. On top of that, questions with explicit negatives sometimes generate invalid responses, because fast responders miss the negative term and therefore provide a response that does not match their opinion (e.g., Dillman et al. 2009).

Taken together, based on survey handbooks and linguistic research, we may assume that negative questions are more difficult to process than their positive counterparts. As no-opinion answers are an important proxy for question complexity (e.g., Deutskens et al. 2004; Kamoen and Holleman 2017), we expect more no-opinion answers for negative questions than for positive ones. We will test this hypothesis in the context of a specific type of survey called a voting advice application (VAA). VAAs are online tools that help users determine which party to vote for in election times. These tools have become increasingly popular in Europe over the past decades, reaching up to 40% of the electorate in countries such as the Netherlands (see Marschall 2014). In a VAA, users express their attitudes to a set of survey questions about political issues. These questions are formulated by a commercial or government-funded VAA developer, in dialogue with the political parties running in the election. Based on the match between the user’s answers and the parties’ issue positions, the tool subsequently provides a personalized voting advice. In the calculation of this advice, no-opinion answers

are excluded (De Graaf 2010; Krouwel et al. 2012). As VAA developers want to base their voting advice on as many VAA questions as possible, this makes an investigation of the effect of question polarity on the proportion of no-opinion answers practically relevant too; it would be problematic if one wording would lead to more no-opinion answers than another wording. This is especially true since several studies have shown that the VAA voting advice has an impact on users' vote choice (e.g., Andreadis and Wall 2014; Wall et al. 2012).

Method

Design and Materials

During the Dutch municipal elections of March 2014, we conducted a real-life field experiment on a VAA developed for the municipality of Utrecht, which is the fourth largest city in the Netherlands with 258,087 inhabitants. In collaboration with VAA developer Kieskompas and all of the 17 political parties running in the elections, four experimental versions of Kieskompas Utrecht were constructed in addition to this benchmark version.¹

In the experimental VAA versions, the polarity of the question was varied for 16 out of 30 questions (see Figures 1 and 2 for an example). These manipulations can be distinguished into two types. A total of 10 questions contained an explicit sentence negation (e.g., 'The municipality *can cut down/cannot cut down* on social work'). The remaining 6 manipulations contained an implicit negative, so a word with negative valence (e.g., The requirement for a building permit of one's own house should *remain to exist /be abolished*). Research shows that language users can easily distinguish between words with positive versus negative valence (Hamilton and Deese 1971), and all the implicit negative terms used in the current research scored high on negative valence in an empirical study (Warriner et al. 2013)². All experimental materials can be found in supplemental materials.

The manipulated questions were distributed across the VAA versions in such a way that each VAA contained an equal number of positive and negative items. All of the Kieskompas Utrecht visitors ($N = 41,505$) were randomly assigned to either

¹ The project description was approved prior to fielding the study by Utrecht University, the Dutch Science Foundation, and the Utrecht City Council. Visitors always entered Kieskompas Utrecht voluntarily, and they could stop filling out the VAA at any point in time.

² Warriner et al. (2013) report valence scores on a scale from 1 (negative) to 9 (positive) for 13,915 English words. Human evaluators assign these scores. Across all reported words, the average valence score is 5.06. The implicit negative terms used in the current research (forbid, ban, stop, abolish, and force) had valence scores between 2.82 and 4.73. For the positive equivalents (allow, decide for yourself, remain to exist, maintain, continue) scores ranged between 5.61 and 6.39.

the standard Kieskompas version, or to one of the experimental versions.³

The screenshot shows a survey interface with a progress bar at the top (1-30, with 5 highlighted). The question is under the heading 'Mobiliteit' and reads: 'De meest vervuilende auto's (ouder dan Diesel Euro 3 en Benzine Euro 0) moeten in de binnenstad worden toegelaten'. Below the question are six response buttons: 'helemaal mee eens', 'mee eens', 'neutraal', 'niet mee eens', 'helemaal niet mee eens', and 'geen mening'. At the bottom, there are social media sharing options for 'Like' (87) and 'Tweeten' (872).

Figure 1 Example of a positively worded question: The cars that are most polluting (cars older than Diesel Euro 3 and Gas Euro 0) should be allowed in the city center. The concise translation of the response categories is: “completely agree, agree, neutral, not agree, completely not agree, no-opinion”.

³ A total of 13 out of the 16 manipulated questions also contained a manipulation of issue framing, operationalized by variation in the heading above the question (left-wing or right-wing). So, in fact, these questions were manipulated following a 2 (question polarity: positive or negative) x 2 (heading: left-wing or right-wing) design. As the effect of question polarity did not interact with the effect of the headings, we decided to report the effect of the headings elsewhere (authors, under review).

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

Mobiliteit

De meest vervuilende auto's (ouder dan Diesel Euro 3 en Benzine Euro 0) moeten uit de binnenstad worden geweerd

helemaal mee eens mee eens neutraal niet mee eens helemaal niet mee eens geen mening

Like 87 Tweeten 872

Figure 2 Example of a negatively worded question: The cars that are most polluting (cars older than Diesel Euro 3 and Gas Euro 0) should be banned from the city center. The concise translation of the response categories is: “completely agree, agree, neutral, not agree, completely not agree, no-opinion”.

Participants

Kieskompas Utrecht was launched on February 18, 2014. Between February 18 and March 19 (Election Day), the tool was visited 41,505 times.⁴ For the purposes of the current study, we focus on those VAA users who were randomly assigned to one of the experimental versions of Kieskompas, which means that VAA users who were assigned to the benchmark version ($N = 7,812$) were excluded. This was necessary because the benchmark version differed in more than one respect from the four experimental versions, as the benchmark version did not contain headings above the question.

In addition, we only took into account those VAA users who were 18 or older (and hence eligible to vote), for whom it took longer than 2 minutes to fill out all 30 statements, and who did not show straight-lining behavior (i.e., reported the same answer to each and every statement). This cleaning method is similar to

⁴ It is impossible to check whether all visitors are unique users, because monitoring IP-addresses would violate VAA users' privacy concerns. Even if IP-addresses would be available, it would be impossible to distinguish unique users based on IP-address because multiple users may access the tool from the same IP-address and the same user may access the tool from various IP-addresses. If the same users filled out the VAA twice (or more), they would again be assigned randomly to a VAA version and receive a VAA consisting of positive and negative questions. We expect that multiple usage might decrease effect sizes, but will not affect the direction of the effects.

the one used in Van de Pol et al. (2014). Cleaning the data led to the exclusion of another 2,581 cases, which means that 31,112 Kieskompas users were included in the analyses.

In our final sample, the male/female division is about equal (50.7% female). The mean age is 37.3 years ($SD = 13.8$). VAA users are fairly highly educated (the median category was higher vocational education or university bachelor), and rather interested in politics (mean of 3.3 on a 5-point scale, $SD = 0.83$). These imbalances with respect to educational level and political interest are very common for samples of VAA users (Marschall 2014).

In order to check the randomization, we compared the experimental versions with respect to age ($F(3, 22207) = 1.58$; $p = .19$), gender ($\chi^2(3) = 2.90$; $p = .41$), educational level ($\chi^2(3) = 11.90$; $p = .85$) and interest in politics ($F(3, 21990) = 0.34$; $p = .80$). As none of these tests showed a difference between conditions, there is no reason to assume that there are *a priori* differences between the VAA users in the experimental conditions.

Measurement and analyses

To analyse the effect of statement polarity on the proportion of no-opinion answers, we constructed a binary variable indicating whether the VAA user provided a no-opinion answer (0) or a substantive answer (1) to each of the 16 manipulated questions. This binary dependent variable was subsequently predicted in the logit multilevel model displayed in Equation 1 below. In this model, $Y_{(jk)}$ indicates whether or not individual j ($j = 1, 2, \dots, 31,112$) gives a substantive answer to question k ($k = 1, 2, \dots, 16$). In the model, two cell means (Searle, 2006) are estimated in the fixed part of the model: one for positive and one for negative question wordings.

To estimate these, dummy variables are created that can be turned on if the observation matches the prescribed type (D_POS or D_NEG). Using these dummies, two logit proportions are estimated (β_1 , and β_2), which may vary between persons (u_{1j} , u_{2j}) and questions (v_{0k})⁵. The model assumes that the proportion of no-opinion answers is nested within items and respondents at the same time. This means that a cross-classified model is in operation (Quené

⁵ Please note that the model implies that there is variance due to the interaction between respondent and item. However, because the dependent variable in the model is binomial, this variance is not estimated as it is fixed if the mean proportions are known. The interaction variance can be approximated applying the formula $p * (1-p)$, in which p represents the estimated proportion of no-opinion answers for positive and negative questions, respectively.

and Van den Bergh 2004; 2008). Please note that while the person-variance is estimated separately for positive and negative wordings, the question variance is estimated only once. This is a constraint of the model. All residuals are normally distributed with an expected value of zero, and a variance of, respectively, S^2_{u1j} , S^2_{u2j} , and S^2_{v0k} .

Equation 1:

$$\text{Logit } Y_{(jk)} = D_POS_{(jk)}(\beta_1 + u_{1j}) + D_Neg_{(jk)}(\beta_2 + u_{2j}) + v_{0k}$$

Results

The first row in Table 1 shows the mean proportion of substantive answers across the 16 manipulated statements. A comparison of positive and negative wordings shows that there is no effect of question polarity on the number of nonsubstantive answers ($\chi^2 = 0.04$; $df = 1$; $p = .95$).

Table 1 Parameter estimates of the multilevel models used for estimating the effect of question polarity ($N_{\text{respondents}} = 31,112$).

	% substantive answers (Logit; SE)		$S^2_{\text{respondents}}$ (SE)		S^2_{items} (SE)
	Positive	Negative	Positive	Negative	Pooled
All items	94.7%	94.7%	2.86	2.67	0.91
($N_{\text{items}} = 16$)	(2.88; 0.24)	(2.88; 0.24)	-0.04	-0.04	-0.32
Explicit negatives	94.2%**	93.9%	2.21	2.01	1.09
($N_{\text{items}} = 10$)	(2.79; 0.33)	(2.74; 0.33)	-0.05	-0.04	-0.49
Implicit negatives	95.4%**	95.9%	2.89	3.23	0.21
($N_{\text{items}} = 6$)	(3.03; 0.19)	(3.16; 0.19)	-0.08	-0.09	-0.12

Note Table 1. For the sake of presentational clarity, the mean answers are given in percentages and in the Logits used for the analysis (between brackets). A higher percentage means that more substantive answers are provided. The variances are only expressed in Logits.

** = $p < .001$

As we did not observe an overall polarity effect, we also explored the effect of question polarity for each of the two types of negatives separately. The second row of Table 1 shows the polarity effect for explicit negatives ($N = 10$). In line with prior expectations, we observed that negative questions generate more no-opinion answers than their positive equivalents ($\chi^2 = 14.00$; $df = 1$; $p < .001$). The size of this effect, however, is tiny relative to both the between-person standard deviation (Cohen's $d = 0.03$) and the between-question standard deviation (Cohen's $d = 0.05$). In absolute terms, the chance of providing a no-opinion answer is about 5% larger for negative questions than for positive ones.

The third row of Table 1 displays the effect of question polarity for the subset of items containing an implicit negative ($N = 6$). Also for this subset of items, an effect of question polarity is observed, albeit in a different direction: contrary to expectation, positive questions yield more no-opinion answers than their negative equivalents ($\chi^2 = 35.5$; $df = 1$; $p < .001$). The size of this effect is small compared to the differences between respondents (Cohen's $d = 0.07$), and substantive but small compared to the differences between items (Cohen's $d = 0.28$). The chance of providing a no-opinion answer is roughly 14% larger for positive questions than for negative ones.

Conclusion

The current research investigated, in the context of an online VAA, whether the proportion of no-opinion answers depends on an important question characteristic: the choice for a positive or a negative statement wording. Across a set of 16 manipulated questions we find no overall effect of question polarity. This is contrary to expectations, because survey handbooks (e.g., Dijkstra and Smit 1999; Dillman et al. 2009; Korzilius 2000) as well as linguistic research (Clark 1976; Kaup et al. 2006) point out that negative questions and their answers are structurally more difficult to comprehend than their positive counterparts are. We did observe polarity effects when analysing two types of negatives separately. For questions including an explicit negation (e.g., not or none), we observed more no-opinion answers for the negative question versions as compared to their positive equivalents. The reverse was true for the set of implicit negatives (e.g., forbid/allow): For these pairs, the positive wording generated more no-opinion answers.

Discussion

We can only speculate about the reasons for the unexpected finding that implicit negatives yield less no-opinion answers than their positive equivalents. One explanation is that implicit negatives are actually easier to process than their positive counterparts. This explanation matches with work on the forbid/allow-asymmetry. In a semantic analysis of forbid and allow questions, it has been shown that the meaning of forbid-questions is really well-defined, as forbidding always indicates "...an act of inserting a barrier, and to a force dynamic pattern which brings about change", whereas allow-questions are more ambiguous, because allowing "may imply causing (removing a barrier) as well as letting (not inserting a barrier)" (Holleman 2000, 186). In line with this work, it has been shown that the answers to a set of forbid-questions are more homogeneous,

and therefore more reliable, than answers to a comparable set of allow-questions (Holleman 2006). Hence, if these results generalize to other contrast pairs, results of the present study can be explained by the fact that implicit negatives have a clearer meaning as compared to positive equivalents.

Alternatively, not (just) the linguistic form, but the match between the linguistic form and the status quo of the question topic may explain our findings. In our experimental materials, the items including an implicit negative (e.g., forbid/allow) always related to situations where the status quo matched the positive wording (e.g., circuses with animals are currently allowed). By contrast, for items including an explicit negation (e.g., may/may not cut down), the negative wording always matched the status quo (e.g., there are no cut-downs on art and culture). This means that we can also rephrase our results such that, irrespective of the linguistic form, wordings representing the current state of affairs generate more no-opinion answers than wordings that represent change. The idea that the appropriateness of a linguistic form in its usage context determines processing complexity, matches with theories from pragmatic linguistics (e.g., Sperber and Wilson's Relevance Theory 1995). According to this work, language users want to make their contribution to a discourse situation as informative as required for the purposes of the exchange. In a political attitude context, wordings that represent change with respect to the status quo are probably more informative than wordings that describe the current state of affairs. This is because most citizens have little knowledge of the exact political issues at stake (Delli Carpinì and Keeter 1996), whereas they do have knowledge about the status quo, as they encounter this status quo in their daily lives (Lupia 1992). To disentangle these two explanations, we propose a future study in which both the question wording (positive, implicit negative, explicit negative) and the status quo (issue X is currently allowed in municipality 1/ and forbidden in municipality 2) are varied. Such an experiment will allow tearing apart the two explanations.

Although we cannot yet explain the current findings, they are clearly relevant for survey and VAA practice. In a VAA context, the answers VAA users give to the political attitude questions directly influence the voting advice (De Graaf 2010; Krouwel et al. 2012). As the voting advice affects vote choices (e.g., Andreadis and Wall 2014), we believe that it is legitimate to conclude that our results are in fact important for VAA developers, even though the statistical size of the effects observed is small. Taking the odds ratio as a standard of comparison, about 1 in 20 no-opinion answers can be avoided if explicit negations matching the status quo are replaced by positive wordings representing change. Hence, if there is a political discussion about whether or not new houses should be built in a certain area, a question wording such as "New houses should be built in area X" should

be preferred over the negative phrase “There should be no new houses built in area X”. In addition, roughly one out of every 7 no-opinion answers can be avoided when positive wordings that match the status quo are replaced by implicit negatives describing change. Hence, if there is a debate about whether taxes on housing should or should not continue to exist, one can better ask respondents to react to the statement “Taxes on housing should be abolished”, rather than “Taxes on housing should remain to exist”.

Moreover, our results are also relevant for the broader context of political attitude surveys and surveys on policy issues, as these surveys include very similar questions to the ones we find in VAAs. Inspect, for example, questions in Eurobarometer (e.g., asking if respondents are pro or con a European economic and monetary union with one single currency, the Euro; Eurobarometer 2015, QA18.1), or popular polls in newspapers and other media (e.g., Do you think Ukraine should become a member of the EU?; <https://www.burgercomite-eu.nl/peiling-maurice-de-hond/>). The only difference between a VAA and these other political attitude survey contexts might be the type of respondents: VAA users may be more motivated to fill out the questionnaire as they are rewarded with a personalized advice (Holleman et al. 2013). Moreover, the VAA users in our sample appeared to be rather highly educated and fairly interested in politics. We know from Krosnick’s work on survey satisficing (e.g., Krosnick 1991) that the more motivated, highly educated and interested in the survey topic the respondent is, the smaller the size of the effect of various wording variations on reported attitudes. Hence, if VAA users are really more motivated (but studies showing otherwise are Baka et al. 2012; Kamoen and Holleman 2017), highly educated and interested, this would imply that polarity effects may be even larger in these other political attitude contexts.

Overall, we conclude that questions about political issues can best be phrased in terms of a change with respect to the status quo, and in doing so, not to shy away from using implicit negatives; using this technique can reduce the proportion of no-opinion answers. Therefore, if a country is currently in the European monetary union, it is better to ask if the country should leave this union rather than ask about staying in.

Funding

This work was supported by the Dutch Science Foundation (NWO), grant number 321-89-003.

References

- Andreadis, I. and M. Wall. 2014. The impact of voting advice applications on vote choice. In: (D. Garzia and S. Marschall, eds.) *Matching voters with parties and candidates: voting advice applications in comparative perspective*. ECPR Press, Colchester, UK, pp. 115–128.
- Baka, A., L. Figgou and V. Triga. 2012. ‘Neither agree, nor disagree’: A critical analysis of the middle answer category in voting advice applications. *International Journal of Electronic Governance* 5(3-4): 244–263. doi: 10.1504/IJEG.2012.051306.
- Chessa, A.G. and B.C. Holleman. 2007. Answering attitudinal questions: modelling the response process underlying contrastive questions. *Applied Cognitive Psychology* 21(2): 203–225. doi 10.1002/acp.1337.
- Clark, H.H. 1976. *Semantics and comprehension*. The Hague, Netherlands, Mouton.
- Clark, H.H. and E.V. Clark. 1977. *Psychology of language: an introduction to psycholinguistics*. Harcourt Brace, New York.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, New York.
- De Graaf, J. 2010. The irresistible rise of Stemwijzer. In: (L. Cedroni and D. Garzia, eds.). *Voting advice applications in Europe: the state of the art*. Scriptaweb, Napoli, Italy. pp. 35–46.
- Delli Carpini, M. X. and S. Keeter. 1996. *What Americans know about politics and why it matters*. Yale University Press, New Haven, CT.
- Deutskens, E., K. De Ruyter, M. Wetzels and P. Oosterveld. 2004. Response rate and response quality of Internet-based surveys: an experimental study. *Marketing Letters* 15(1): 21–36.
- Dijkstra, W. and J. Smit. 1999. *Onderzoek met vragenlijsten. Een praktische handleiding*. VU Uitgeverij, Amsterdam, Netherlands.
- Dillman D., J.D. Smith and J.M. Christian. 2009. *Internet, mail, and mixed mode surveys: the tailored design method*. Wiley, Hoboken, NJ.

Friborg, O., M. Martinussen and J.H. Rosenvinge. 2006. Likert-based versus semantic differential-based scorings of positive psychological constructs: a psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences* 40(5): 873–884. <https://doi.org/10.1016/j.paid.2005.08.015>.

Hamilton, H.W. and J. Deese. 1971. Does linguistic marking have a psychological correlate? *Journal of Verbal Learning and Verbal Behavior* 10(6): 707–714. [https://doi.org/10.1016/S0022-5371\(71\)80079-8](https://doi.org/10.1016/S0022-5371(71)80079-8).

Holleman, B.C. 2000. *The forbid/allow asymmetry. On the cognitive mechanisms underlying wording effects in surveys*. Rodopi, Amsterdam.

Holleman, B.C. 2006. The meanings of ‘yes’ and ‘no’. An explanation for the forbid/allow asymmetry. *Quality and Quantity* 40(1): 10–38. doi: 10.1007/s11135-005-4479-6.

Holleman, B.C., N. Kamoen and C.H. De Vreese. 2013. Stemadvies via Internet: antwoorden, attitudes en stemintenties. *Tijdschrift voor Taalbeheersing* 35(1): 25–46.

Holleman, B.C., N. Kamoen, A. Krouwel, J. Van de Pol and C.H. De Vreese. 2016. Positive vs. negative: the impact of question polarity in voting advice applications. *Plos One* 11(10). <http://dx.doi.org/10.1371/journal.pone.0164184>.

Hoosain, R. 1973. The processing of negation. *Journal of Verbal Learning and Verbal Behaviour* 12(6): 618–626.

Horn, L.R. 1989. *A natural history of negation*. Chicago University Press, Chicago.

Kamoen, N. and B.C. Holleman. 2017. I don’t get it. Response difficulties in answering political attitude statements in voting advice applications. *Survey Research*

Methods 11(2): 125–140. doi:10.18148/srm/2017.v11i2.6728.

Kamoen, N., B.C. Holleman and H. Van den Bergh. 2013. [Positive, negative, and bipolar questions: the effect of question polarity on ratings of text readability](#). *Survey Research Methods* 7(3): 181–189.

Kaup, B., J. Ludtke and R.A. Zwaan. 2006. Processing negated sentences with

contradictory predicates: is a door that is not open mentally closed? *Journal of Pragmatics* 38(7): 1033–1050. doi:10.1016/j.pragma.2005.09.012.

Korzilius, H. 2000. *De kern van het survey-onderzoek*. Van Gorcum, Assen, Netherlands.

Krosnick, J.A. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5(3): 213–236.

Krosnick, J. A. and S. Presser. 2010. Questionnaire design. In: (D. Wright and P.V. Marsden, eds.) *Handbook of survey research* (2nd ed.). Emerald Group, West Yorkshire, England, pp. 263–314.

Krouwel, A.P.M., T. Vitiello and M. Wall. 2012. The practicalities issuing vote advice: a new methodology for profiling and matching. *International Journal of Electronic Governance* 5(3-4): 223–243.

Lupia, A. 1992. Busy voters, agenda control, and the power of information. *The American Political Science Review* 86(2): 390–403. doi: 10.2307/1964228.

Marschall, S. 2014. Profiling users. In: (D. Garzia and S. Marschall, eds.) *Matching voters with parties and candidates: voting advice applications in comparative perspective*. ECPR Press, Colchester, UK. pp. 93–106.

Quené, H. and H. Van den Bergh. 2004. On multilevel modeling of data from repeated measures designs: a tutorial. *Speech Communication* 43(1-2):103–121. doi:10.1016/j.specom.2004.02.004

Quené, H. and H. Van den Bergh 2008. Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59(4): 413–442. doi:10.1016/j.jml.2008.02.002.

Rugg, D. 1941. Experiments in wording questions II. *Public Opinion Quarterly* 5(1): 91–92. <https://doi.org/10.1086/265467>.

Saris, W., M. Revilla, J.A. Krosnick and E.M. Shaeffer. 2010. Comparing questions with agree/disagree response options to questions with construct-specific response options. *Survey Research Methods* 4(1): 61–79.

Schuman, H. and S. Presser. 1981/1996. *Questions and answers in attitude surveys. experiments on form, wording and context*. Academic Press, London, England.

- Searle, S.R. 2006. *Linear models for unbalanced data*. Wiley, New York.
- Sherman, M.A. 1973. Bound to be easier? The negative prefix and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior* 12(1): 76–84. [https://doi.org/10.1016/S0022-5371\(73\)80062-3](https://doi.org/10.1016/S0022-5371(73)80062-3).
- Sperber, D. and D. Wilson. 1995. *Relevance: communication and cognition* (2nd ed.). Blackwell Publishers, Oxford/Cambridge.
- Sudman, S. and N.M. Bradburn. 1982. *Asking questions. a practical guide to questionnaire design*. Jossey-Bass Publishers, San Francisco, CA.
- Swain, S.D, D. Weathers and R.W. Niedrich. 2008. Assessing three sources of misresponse to reversed Likert items. *Journal of Marketing Research* 45(1): 116–131. <http://dx.doi.org/10.1509/jmkr.45.1.116>.
- Van de Pol, J., B.C. Holleman, N. Kamoen, A. Krouwel and C.H. De Vreese. 2014. Beyond young, higher educated males: a typology of VAA users. *Journal of Information Technology and Politics* 11(4): 397–411. doi: 10.1080/19331681.2014.958794.
- Wall, M., A.P.M. Krouwel and T. Vitiello. 2012. Do voters follow the recommendations of voter advice application websites? A study of the effect of kie.sKompas.nl on its users' vote choices in the 2010 Dutch legislative elections. *Party Politics* 20(3): 416–428. doi: 10.1177/1354068811436054.
- Warriner, A.B., V. Kuperman and M. Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* 45(4): 1191–1207. doi: 10.3758/s13428-012-0314-x.
- Weisberg, H.F. 2005. *The total survey error approach: a guide to the new science of survey research*. The University of Chicago Press, Chicago.

Figures, Tables, and Supplementary Materials

Figure 1 Example of a positively worded question: The cars that are most polluting (cars older than Diesel Euro 3 and Gas Euro 0) should be allowed in the city center. The concise translation of the response categories is: “completely agree, agree, neutral, not agree, completely not agree, no-opinion”.

Download: <http://app.scholasticahq.com/api/v1/attachments/10028/download>

Figure 2 Example of a negatively worded question: The cars that are most polluting (cars older than Diesel Euro 3 and Gas Euro 0) should be banned from the city center. The concise translation of the response categories is: “completely agree, agree, neutral, not agree, completely not agree, no-opinion”.

Download: <http://app.scholasticahq.com/api/v1/attachments/10029/download>

Table 1 Parameter estimates of the multilevel models used for estimating the effect of question polarity ($N_{\text{respondents}} = 31,112$).

Download: <http://app.scholasticahq.com/api/v1/attachments/10030/download>

Appendix 1

Download: <http://app.scholasticahq.com/api/v1/attachments/10032/download>
